Reprinted from Female Genital Cancer. Gusberg, S.B., Shingleton, H.M., Deppe, G. (eds.). New York, Churchill Livingstone, 1988, pp. 21-31

2

The Epidemiologic Method of Research

Robert N. Hoover
Joseph F. Fraumeni, Jr.

The study of human populations in an attempt to identify the causes of disease has been a fascination of physicians and laypersons alike for some time. It was raised by Hippocrates in his treatise, On Airs, Waters, and Places[1]:

> Whoever wishes to investigate medicine properly should proceed thus: in the first place to consider the seasons of the year . . . then the winds, the hot and the cold, . . . one should consider most attentively the waters which the inhabitants use . . . and the ground . . . and the mode in the which the inhabitants live, and what are their pursuits, whether they are fond of drinking and eating to excess, and given to indolence, or are fond of exercise and labor.

However, most of the meaningful work in this area has only occurred during the last 150 to 200 years. For most of this more recent time, the fascination with study of diseases in human populations had focused primarily on the epidemics of the major scourges of infectious disease that had plagued mankind for centuries. With the elucidation of the causes of these diseases, and the opportunities for prevention that have flowed from these investigations, both the quantity and quality of life have improved to the point where the chronic diseases have become the major scourges that attract the interest and concern of the clinician and layperson. Thus, it has only been very recently, primarily the past 40 years, that studies of disease in human populations have focused on malignancy. While important etiologic insights have come from these investigations for a variety of cancer sites (e.g., cigarette smoking and lung cancer), the organ system that has received the most attention, and for which the number and variety of risk factors identified had been most numerous, is the female reproductive tract.

From the earliest observations of the rarity of cancer of the uterine cervix among nuns,[2] interest in risk factors for this organ has focused on sexual practices and related suspect infectious agents. Current enthusiasm over such agents, particularly papillomaviruses, keep this a hot topic for both research and debate in gynecology.[3] For endometrial cancer, early observations concerning the role of obesity, polycystic ovarian disease, and other risk factors, raised the likelihood of an estrogenic etiology.[4] The subsequent epidemic of exogenous estrogen-induced endometrial cancer, controversies over the indications for such treatment, the influence of added progestins, and risk-benefit questions, have all brought the epidemiology of this disease into the everyday discussions of gynecologists. Longstanding anxiety over possible carcinogenic effects of the oral contraceptive, supported by studies suggesting such a role for cancers of the cervix,[5] and muted by the observations of protection for endometrial and ovarian cancer,[6] has also become a part of the daily practice of obstetrics and gynecology. This same clinical community has also had to deal with suggestions that things as personal as the patient's use of body talc, and the surgeon's use of talc on his gloves may in some way relate to the risk of subsequent ovarian cancer.[7]

In the face of these observations and many others, the gynecologist, and particularly the gynecologist interested in malignancy, is confronted with the frequent task of interpretation of epidemiologic findings, with opportunities for scientific collaboration with epidemiologists, and with prospects for making etiologic observations at the bedside himself. For these reasons, it is probably more important for this specialty to understand the principles and methods of epidemiologic investigation than it is for any other clinical specialty.

DEFINITION

The root of the word *epidemiology* is *epidemic*. This is entirely appropriate, since the methods of the discipline are primarily those developed for the investigation of classic epidemics. On a practical level, also, the concept of excessive frequency of disease that is implied by the word epidemic is central to all epidemiologic investigations. Clearly, the goal of an epidemiologic investigation is to establish why, or if, a particular disease is excessive in a specific population group. However, since the word in its general usage seems to engender very subjective and differing opinions of how excessive the frequency must be to qualify as epidemic, thinking about epidemiology in these terms can often distract from an understanding of the principles involved. For example, in utero exposure to diethylstilbestrol (DES) would be unlikely to be considered responsible for an epidemic of gynecological malignancies when it has been responsible for less than one-tenth of 1 percent of all gynecologic malignancies diagnosed in the United States over the past 20 years. However, in terms of vaginal adenocarcinoma among young women, the epidemic nature of the same observed number of cases is undeniable. Likewise, the marked decline in invasive cancers of the uterine cervix over the past 30 years would seem to be the antithesis of an epidemic. Using the same data, however, the rate of this disease among women aged 45 to 49 would be considered to be clearly epidemic compared with the rate among women aged 15 to 19.

Because of these semantic nuances, it is probably more useful not to think of epidemiology in terms of epidemics but rather in terms of its most popular textbook definition, that is, "the study of the distribution and determinants of disease frequency in human populations."[8] Although brief, this definition uses a number of key words:

1. *Epidemiology* refers to humans, a feature distinguishing it from a variety of other disciplines in cancer research.
2. It entails the study of populations. This distinguishes it from other disciplines, such as clinical research, which also study disease processes in human beings, but at the individual or case-series level. By contrast, epidemiology is the study of groups of people, groups definable in a variety of different ways that would permit them to be characterized as a population.
3. The term *frequency* denotes the quantitative orientation of the discipline. Epidemiology is a numerate science, based on the principle that if you can't count it, you don't understand it. Thus, rather than seeking solely qualitative differences, the epidemiologic method attempts to quantify the risks of disease attributable to various causes.
4. The terms *distribution* and *determinants* describe the two major approaches of the epidemiologic method. Descriptive studies examine the distribution of disease and are usually employed to generate etiologic hypotheses, while analytic studies are used mainly to test hypotheses and identify the determinants of disease. A primary objective of epidemiology is to identify and quantify relationships between exposure to environmental agents and deleterious health effects. These associations may lead to causal inferences, which in turn provide the basis for instituting preventive measures for various diseases.

EPIDEMIOLOGIC INVESTIGATIONS

Descriptive and Correlational Studies

Descriptive (or demographic) studies are concerned with identifying the distribution or patterns of disease in populations.[8] It is a basic tenet of epidemiology that diseases, including cancer, do not occur randomly, but fluctuate according to factors such as age, sex, race, time, and geographic location. The use of rates as measures of disease frequency is fundamental in describing patterns of cancer among these population groups. Prevalence, incidence, and mortality rates of cancer define the levels of risk prevailing in different populations and permit comparisons between groups. Descriptive surveys of cancer occurrence have been valuable in stimulating etiologic hypotheses and in providing direction for analytical studies, which are then necessary to establish whether risks are associated with particular exposures.[9] Thus, important leads to etiology have come from population-based cancer surveys, which have demonstrated substantial international variations in cancer incidence, shifts in risk among migrant populations, changes in risk over time, and geographic peculiarities from mapping cancer mortality at the county level.[10]

Descriptive studies may use the correlational (or ecological) approach, in which the rate of disease in a population is compared with the spatial or temporal distribution of suspected risk factors.[9] This type of study may be particularly helpful in developing or refining hypotheses about carcinogenic risks but falls short of establishing causal relationships.

Correlational studies have the advantage of being much less expensive and time consuming than analytic investigations because they often use mass statistics previously collected for another purpose.[11] The primary weakness of such studies, as with descriptive studies generally, is that data are collected on populations rather than individuals. In other words, the rate of disease and the prevalence of exposures to variables of interest are known for various population groups, but information on the exposure status of persons who have the disease and those who do not within each population is lacking.

Thus, one cannot infer from the correlations at the population level that the exposure of concern is associated with the risk of developing disease within each population.[11] For example, in early surveys of lung cancer, the international variation in mortality rates and temporal increases among males appeared consistent with the reported patterns of cigarette smoking, but these correlations by themselves may have been circumstantial rather than causal, since a variety of other exposures (e.g., occupational hazards, air pollution) also varied concomitantly with the patterns of lung cancer. It took the analytic studies that pursued these leads to establish the cause-and-effect relationships between smoking and lung cancer.

Correlational studies may also provide supporting evidence in evaluating relationships detected by analytic studies or laboratory data. This is illustrated by the more recent temporal increase in lung cancer among females, who have lagged about 20 years behind males in their adoption of smoking habits. Because correlational studies deal with aggregate exposures and disease occurrence at the population level, they are often also seriously limited by the imprecise measurements of exposure and the many potentially confounding variables. A relevant example in regard to these points concerns the relationship between menopausal estrogen use and the risk of endometrial cancer. One study evaluating the time trends in endometrial cancer from the late 1940s to the early 1970s noted a lack of any appreciable increase over this period,[12] leading some workers to conclude that this exonerated any recently introduced risk factors, including the use of menopausal estrogens. A subsequent evaluation of time trends during the 1970s reached the opposite conclusion.[13] Certainly the analytic studies of these issues have solidly supported the conclusions of the latter investigation.

Many reasons have been suggested for the failure of the first descriptive study to note a trend that could have pointed to newly introduced risk factors, the most prominent being the relatively small proportion of women exposed in the period covered and, perhaps more importantly, the rising rates of hysterectomy over this same time period, which tended to reduce the real number of women at risk of developing this tumor. Whatever the explanation, the dangers of the shift to causal interpretations on the basis of descriptive studies should be clear.

Analytical Studies

In order to test etiologic hypotheses and to identify and quantify carcinogenic risks to humans, it is necessary to conduct analytic epidemiologic studies.[8,9,11] These studies are the principal means of determining the human health hazards of specific environmental exposures and agents. In contrast to descriptive surveys, data are obtained on disease occurrence and putative risk factors for specific individuals, using mainly the case-control or cohort method. Thus, by grouping exposed individuals and comparing them to those unexposed, after controlling for all other relevant variables, the risk of disease associated with exposure can be estimated. While it is important to avoid imposing unnecessary constraints on epidemiologic investigation, some methodologic guidelines should be considered in designing a study. In particular, the study groups should be sufficiently large, and the time intervals between initial exposure and tumor onset sufficiently long, to identify the lowest excess risk considered important to detect. Reliable and valid estimates of exposure should be sought, with quantitative measurements to permit dose-response evaluations. Studies should be designed in a manner that minimizes potential sources of bias, and permits detection and control of confounding variables.

CASE-CONTROL STUDIES

Case-control studies start by identifying persons with a particular disease (cases) and a group of similar persons without the disease (controls). Information on past exposure to known or suspected risk factors is then collected from interviews, questionnaires, medical records, occupational logs, or other sources. The frequency of a particular exposure among the cases is compared with that in the control group, after making appropriate adjustments for other relevant differences between the two groups. If the proportion of cases with a certain exposure is significantly greater than that of the controls, an association between exposure and disease may be indicated. The case-control approach is especially well suited in studying relatively rare conditions, such as most cancers, in which the putative exposure is common in the general population (e.g., menopausal estrogens and endometrial cancer), or when the exposure is rare but accounts for a large portion of a particular cancer (e.g., DES and vaginal adenocarcinoma).[14]

COHORT STUDIES

Cohort studies begin by identifying a group of individuals with a particular exposure and a similar group of unexposed persons and following both groups over time to determine subsequent health outcomes. The rates of disease in the exposed and unexposed groups are then compared. Information on disease frequency and other factors may be identified from medical records, occupational records, physical examinations, interviews, questionnaires, or death certificates. An association between exposure and disease may be indicated if the rates of disease are greater in the exposed group than in the unexposed group. These investigations may be based on current exposure and future health outcomes (prospective cohort study), but more commonly they use past exposure information and disease occurrence (retrospective cohort study). Instead of an unexposed comparison group, general population mortality or incidence rates (specific for age, sex, race, and calendar time) are often used to determine the expected number of cases of disease. This method assumes that in the absence of specific exposure the study group would have had the same probability of developing the disease as the general population, but differences in ethnic, socioeconomic, and other variables must be considered in evaluating the validity of this assumption. The cohort approach is used mainly when it is possible to evaluate heavy exposures in clearly defined subgroups of the population. Thus, it has been especially helpful in assessing the carcinogenic risk from occupational hazards or medical exposures, including radiation and certain drugs (e.g., the risk of leukemia following treatment of ovarian cancer patients with alkylating agents).

Both the case-control and cohort methods are characterized as having certain strengths and weaknesses, although they complement each other in testing specific etiologic hypotheses. Case-control studies provide (1) a more efficient means of studying rare diseases, with fewer individuals needed for study as compared with the cohort approach; (2) a shorter time period for study completion and generally lower costs as compared with the cohort method; an opportunity to evaluate simultaneously several causal hypotheses as well as interactions (the extent and manner in which two or more risk factors modify the strength of one another); and (3) a capacity to evaluate the effects of common exposures as well as those rare exposures that may account for a large proportion of the cases. By contrast, the case-control approach has some problems in (1) directly estimating the risk associated with a particular exposure; (2) reducing certain biases (e.g., selection, historical recall) that affect the comparability of cases and controls; and (3) providing detailed and precise information on exposures occurring in the past.[8] By definition, such investigations can only evaluate one disease or outcome at a time.

The advantages of cohort studies are their capacity (1) to estimate directly the risks attributed to a particular exposure, since incidence or mortality from disease is actually being measured; to reduce subjective biases by obtaining information before the disease develops; (2) to determine associations between a particular exposure and multiple health outcomes; and (3) to evaluate temporal relationships such as latency period and duration of effect.

The most obvious temporal relationship that can be established by the cohort method, and in some circumstances established only by this method, is the timing between the exposure of interest and the development of the disease in question. From the first observations of excesses of indices of herpes virus infection among patients with cervical cancer, concerns were raised that such infection might have actually followed the development of the disease, and that the uterine cervix in the first stages of malignancy might, in fact be in some way particularly vulnerable to infection by these viruses.

Virtually the only way to assess these critical questions was to draw bloods from a large population of women for the evaluation of antibodies and then follow them over time to determine the frequency of cervical cancer development in those with and without evidence of prior infection. The first such major investigations designed to address this issue have failed to produce support for the hypothesis that there is a relationship between prior infection and the development of this malignancy.[15] However, cohort studies are usually expensive and complex undertakings. They require (1) large numbers of exposed individuals, particularly when relatively rare events as in the case of most cancers are being investigated; (2) long periods of follow-up to accommodate the latency period for chronic diseases such as cancer; and (3) special handling of problems associated with persons lost to follow-up and with biased estimates of risk as from the healthy worker effect of occupational studies.[8]

INTERVENTION STUDIES

Also referred to as experimental studies,[16] intervention studies represent a third strategy of analytic epidemiology that is especially useful in confirming causal relationships suggested by case-control or cohort studies. This approach may be applied in programs designed, for example, to reduce cigarette smoking and alcohol intake, modify diet, control occupational pollutants, or evaluate candidate preventatives (e.g., vitamin A supplements, the addition of progestins to treatment regimens for menopausal symptoms). Ethical considerations are obviously critical when developing this approach and, after intervention, the statistical procedures resemble those employed for cohort studies.

EPIDEMIOLOGICAL MEASURES

Rates

If epidemiology is a numerate science, the measurement of frequency of disease should be a central feature of the discipline. The basic measure of frequency that has epidemiologic value is the rate. Within this context, this refers to an enumeration of the number of diseased persons expressed per unit size of the population in which these cases were observed. The addition of the element of time to this expression (e.g., time at which cases were observed, or period of time during which they developed) is a key feature, making any rate epidemiologically useful. A typical example in the area of cancer is that the overall incidence rate of cancer in the entire population is estimated to be approximately 325 cases per 100,000 population per year.[17] As can be seen, the expression of disease in this manner permits comparison of the rate of disease in one population with that in another, taking into account the likely prospect that the populations were of different size or were observed over differing periods of time. A variety of mechanisms also exist to adjust these rates for a variety of factors, including age, race, and social class.[18] Such adjustments remove the influence on disease rate of differences among the populations being compared with respect to these variables.

A variety of different types of rates can be developed for measurement of different disease states. For malignancy, the most common rates used are prevalence rates (numbers of cases in existence at a particular point in time divided by the number of people in the population in which they exist), mortality rates (number of deaths due to a particular malignancy over a specified period of time divided by the population under surveillance for these deaths over the same period of time), and incidence rates (number of new cases of a malignancy occurring over a specified period of time divided by the population under surveillance for such cases over the same period of time). Each type of rate has its own particular uses and its own particular drawbacks. For purposes of doing etiologic research, incidence rates are generally preferable. Both prevalence rates and mortality rates reflect the influence of prognosis as well as factors that lead to the occurrence of a disease. Prevalent cases (all those alive in the population at a particular point in time) tend to be heavily weighted with long-term survivors. By contrast, mortality rates clearly reflect the experience of those cases with the poorest prognosis. Incidence rates, reflecting as they do all of the cases of the disease that occur, are representative of the true spectrum of disease. In addition, incidence rates generally measure the frequency of disease at a point much closer in time to etiologic influences than do the other measures.

Measures of Association

The key ingredient in the epidemiologic method is the comparison of attributes between different populations. The question then arises as to how these attributes are to be compared. Considerable interest on the part of clinical and laboratory investigators is focused on tests of the statistical significance of differences between groups. While this assessment of the likely role of chance in producing any differences observed is important, it does not measure the magnitude of the differences and thus the strength of any observed relationship between an exposure and a disease. The main measures of differences between populations utilized by epidemiologists are based on rates of disease. Specifically, two measures are most prominently used. The relative risk RR is the disease rate in the exposed divided by the disease rate in the referent (usually nonexposed) population. This measure gives an estimate of the relative difference in disease risk between the two populations. Thus, a RR value of 2.0 would indicate that the exposed group has twice the risk of the unexposed group (or a 100 percent increase in risk). The other measure of association is the risk difference. As implied, this estimate results from the subtraction of the rate among the unexposed from that among the exposed. This difference in risk is also frequently referred to as the attributable risk, AR. The implication of this terminology is that if the relationship observed is causal, then the difference between the risks is the amount of disease produced among the exposed that is attributable to that exposure.[8]

These two measures have somewhat different uses. The measure of relative risk is usually assumed to be an indicator of how strongly, and thus how likely, an exposure is related to a disease. Thus, the magnitude of the RR value is used as an indicator of how likely the relationship is to be causal. The difference between risks is also influenced by the magnitude of the difference between the exposed and unexposed but is also influenced by the rate of the disease in the absence of exposure. Thus, for a very rare condition, the relative difference between the rates in the exposed and the unexposed can be substantial, but the actual number of cases produced among the exposed could be quite small, owing to the rarity of the disease itself.

A recent follow-up study of 1-year survivors of ovarian cancer from five different randomized trials measured the incidence rate of acute non-lymphocytic leukemia and preleukemia among women treated with no chemotherapy, those treated with cyclophosphamide, and those treated with melphalan.[19] The incidence rates were 0.18, 3.21, and 11.46 cases per 1,000 women per year, respectively. The RR of these conditions compared with those receiving no chemotherapy was 18-fold for women taking cyclophosphamide and 64-fold for those taking melphalan. The RR value for women taking melphalan versus those taking cyclophosphamide was 3.6. The size of the relative excesses for women taking chemotherapy versus those receiving none makes it extremely unlikely that the excess risks are due to anything other than the drugs themselves. On the other hand, while the RR value of 3.6 for melphalan versus cyclophosphamide suggests that there may be a differential leukemogenicity for one drug versus another, the strength of the association is much less than for any chemotherapy versus none and thus requires a more cautious interpretation.

While the RR values are very high for these two alkylating agents as compared with no chemotherapy, the differences between the absolute risks are not very great. Thus, the attributable risk for these conditions among those treated with cyclophosphamide is approximately 3 per 1,000 per year and for melphalan is approximately 11 per 1,000 per year. Given all the other problems confronting ovarian cancer patients, including competing causes of mortality, these excesses probably do not represent a major public health problem. This also explains how it could be difficult for an individual clinician or even a large group practice to notice important etiologic observations, such as these differences in risks. If one were following as many as 100 patients who had been treated with one drug and 100 patients who had been treated with the other, even this 3.5-fold excess risk would result in a differential of only approximately one case of leukemia per year.

One other measure that is sometimes used in epidemiologic investigations is an estimate of the amount of disease attributable to a particular exposure not just among the exposed but in a population that has both exposed and unexposed individuals. This measure would thus reflect the amount of disease that would exist in some definable population if the exposure were removed. This measure is referred to as the population attributable risk or, when expressed as a proportion of the total disease in the population, as the etiologic fraction. This measure is calculated by subtracting the rate among the unexposed from the rate that exists in the total population of interest. It can be seen that the magnitude of this particular estimate relates not only to the magnitude of the relative difference between the exposed and unexposed, and to the level of the disease among the unexposed but to the prevalence of the exposure of interest in the particular population being addressed as well.

Using the example outlined above, even though the relative risk for alkylating agents versus none is very large, if exposure to all alkylating agents were removed, it would have very little impact on the total leukemia rate in the general population, since very few persons in the general population are being exposed to these drugs. By contrast, among the clinical population investigated in this particular study, the overall rate of these leukemic conditions was 2.35 per 1,000 patients per year. Subtracting the rate among those not treated with chemotherapy (0.18 per 1,000 per year) from this yields a population attributable risk of 2.17 cases per 1,000 women per year or, expressed as an proportion of the total observed rate, an etiologic fraction of 92 percent. This high proportion of leukemic disorders among ovarian cancer patients attributable to the use of alkylating agents is a result not only of the high relative risk associated with the use of alkylating agents but of a high prevalence of exposure to these agents in this population of patients.

All these measures have been illustrated in the context of a cohort investigation, within which these rates and risks can be directly measured. In case-control investigations, there are no exposed and unexposed populations per se, and, thus, no ability to calculate rates of disease and relative and attributable risks directly. However, over the years, reliable and reasonable procedures for estimating these parameters within the context of a case-control study have been developed and have become the preferred measures of association in these studies. Relative risks is estimated by the odds ratio, also referred to as the relative odds. By making some assumptions about the representativeness of the exposure among cases and controls with respect to a population to which inferences are to be made, estimates of the differences in risks can be attempted as well.

BIAS AND CONFOUNDING

Epidemiology is primarily an observational rather than an experimental science. Thus, many of the concerns mitigated by a randomization process and a controlled experimental environment have to be addressed specifically by the epidemiologist in the development of the protocol and study procedures, as well as in the analyses and interpretation of data. The major concern is over sources of bias or differential handling of the groups being compared. The result could be the acquisition of groups that are not comparable to each other, and/or the acquisition of data from these groups in a noncomparable manner. These circumstances can arise as a result of actions by the investigator, by the study participants, or by other persons or forces.

Covering all the possibilities for bias is beyond the scope of this review. However, some examples may be illustrative. Two of the more common ways of achieving noncomparable groups are through selection bias and response bias. For example, all the cases of a disease hospitalized at a specialized referral center could be selected for comparison with a random sample of the general population of the area. In this circumstance, any characteristics of the cases that differed from the controls might have nothing to do with the disease but rather could reflect characteristics that would lead to referral to this one institution. Likewise, a differential response rate to a questionnaire (e.g., 90 percent for cases and 60 percent for the controls) could lead to subsets (respondents) of cases and controls that are different from each other, while a comparison of the entire targeted series of cases and controls might yield no such differences.

Bias in the acquisition of data can also come about by a variety of means. One of the most discussed is recall bias or the tendency of cases to recall past exposures more completely because they have been searching their memories for possible reasons for their illness. The tendency of mothers of children with congenital heart defects to report taking virtually every type of medication more frequently during pregnancy than mothers of children without congenital defects is thought to reflect this phenomenon.[20] Similar bias in collection of information concerning cases versus controls can be introduced by others as well. The widely recognized practice of clinicians to record their patients' prior use of exogenous estrogens in a chart only if it would have relevance to a specific problem being worked up leads to a much more complete reporting of such exposure for endometrial cancer patients if chart information from the diagnostic workup is included. To avoid this bias, chart abstraction studies of this issue have had to exclude information recorded in the chart for some arbitrary period of time just prior to diagnosis for the cases and the date of case-matched diagnosis for the controls. This approach undoubtedly results in a failure to identify some cases and controls has having had prior exposure, but this procedure ensures that this will be just as likely for the cases as it will for the controls, these eliminating a biased recording of exposures.

An issue that has received considerable attention in epidemiology, and that will consider to do so, is that of confounding. Confounding of an association between a factor and a disease potentially exists when some other risk factor for the disease (in the context of the study) is also related to the factor under study. This correlation of a another variable with both the exposure and the disease under study can act to produce false associations between the exposure and the disease or to obscure a relationship when one actually exists. A common example is the confounding influence of age. Age is related quite strongly and directly with most malignancies. If this is not accommodated in the design of a case-control study, and a simple random sample of controls is chosen, they will almost always be, as a group, much younger than the cases. If this disparity is also ignored in the analysis, and the crude characteristics of the case group are compared directly with those of the controls, any variable associated with older age will falsely appear to be related to the risk of cancer.

While some confounding relationships are easily recognizable, others can be quite subtle. In a recent study of invasive cervical cancer,[5] the crude estimate of the RR value between this disease and prior long-term use of oral contraceptives was 0.9. However, there was concern that this relationship could be confounded by the time interval between inclusion in the study (diagnosis for the cases) and a prior Papanicolaou (Pap) smear. Since this was invasive cancer, as one might expect, the cases had a much longer time interval between their diagnostic workup and their last previous Pap smear than did the controls. In addition, this same time interval was also related to oral contraceptive use. Oral contraceptive users tended to participate in more regular Pap smear practices and thereby had a much shorter interval between study inclusion and previous Pap smear than did nonusers. Thus, it was believed that a biologic relationship between oral contraceptive use and invasive cervical cancer risk could be obscured by the tendency for oral contraceptive users to be screened more frequently. Indeed, this appeared to be the case. When the relationship was controlled for interval since last Pap smear, the RR value for long-term oral contraceptive use was noted to be 1.8.

STRENGTHS AND LIMITATIONS OF EPIDEMIOLOGY

Strengths

In contrast to studies in other biologic systems, epidemiology directly evaluates the experience of human populations and their response (risk of disease) to various environmental exposures and host factors. Thus, it is often possible to evaluate the consequences of an environmental exposure in the precise manner in which it occurs and will continue to occur in human populations. This includes such important considerations as dose, route of exposure, and concomitant exposures to other exogenous and endogenous factors. Through epidemiologic studies, human cancer has been linked to a number of lifestyle and other environmental hazards, including tobacco products and alcohol, ultraviolet (UV) and ionizing radiation, certain occupational and medicinal chemicals, dietary factors, and some infectious agents.[10,21,22] Epidemiology has played a central role in determining carcinogenic exposures, and it has complemented studies in laboratory animals in clarifying the carcinogenic potential of specific agents.[23] Another strength of the epidemiologic approach is its ability to provide insights into the mechanisms of human carcinogenesis. Thus, epidemiologic observations have complemented experimental evidence that carcinogenesis is a multistage process and that many cancers may result from the cumulative effect of environmental factors and host susceptibility states that accelerate or retard the transition rates at various stages of carcinogenesis.

Limitations

Although epidemiology is the only means of directly assessing the carcinogenic risks of environmental agents in humans, the method has several limitations that are difficult to overcome.[8] One problem is that evidence of an environmental hazard is usually obtained from persons with high or intermediate levels of exposure. Just as for studies in laboratory animals, detecting causal relationships at low exposure levels is difficult, since the observed associations with disease are usually less pronounced and may have alternative explanations, including those related to chance, errors, biases, or confounding variables. To provide a valid basis for risk estimates, large numbers of human subjects are often needed, especially if the exposure is low or rare or if the excess risk is small compared with that of the baseline incidence rate.

Another obstacle to epidemiology is the long latency period between exposure and the development of cancer. This complicates the detection of relationships and makes it impossible to identify the carcinogenic risks to humans of agents newly introduced into the environment. Another common problem in epidemiology is that of exposure assessment. Often the specific exposure of interest cannot be measured directly, so that surrogate measures must be used (e.g., occupation, place of residence).Since exposure data are usually derived from historical records generated for other purposes or from the recollections of subjects, opportunities for either random or biased misclassification of exposure are frequently encountered. In addition, appropriate study groups are often simply unavailable or inaccessible. Furthermore, it may be difficult to implicate specific carcinogens when the environmental hazards involve complex exposures to a variety of agents, the effects of which are difficult to disentangle. Still another difficulty is the inability of epidemiologic studies to adjust for unknown risk factors, since control can be introduced only when the risk factors are already recognized.

Thus, when a particular factor is related to exposure and disease outcome, it may be confounded and give the appearance of an association when in fact none exists, or it may inflate or decrease the magnitude of an association. In view of these difficulties, it is not surprising that epidemiologic data exist for only a small proportion of the many chemicals that have been shown to be carcinogenic in laboratory animals.

BIOCHEMICAL EPIDEMIOLOGY

It seems likely that some limitations of cancer epidemiology may be overcome by incorporating laboratory methods in analytic investigations. This has been a valuable routine practice in infectious disease epidemiology for the past century. This approach, sometimes called biochemical or molecular epidemiology,[15,24] has only recently been developed in cancer epidemiology. There is current enthusiasm for these kinds of investigations, since they merge the strengths of observational human studies with newly developed experimental probes to derive information that could not be developed by epidemiology or laboratory study alone. The laboratory aspect may make it possible to define past exposures and subclinical or preclinical response to initiators, promoters, and inhibitors of carcinogenesis, or to evaluate host-environmental interactions. There is special interest in using this technique to clarify carcinogenic risks associated with nutritional influences or specific environmental agents that can be detected in tissues or body fluids.

Opportunities are also available to assess specific host factors that influence susceptibility to carcinogenesis, including endocrine parameters, immunocompetence, and genetic markers. Techniques are being refined to detect and quantify particular carcinogens or their metabolites in tissues or body fluids through chemical analyses, mutagenesis assays, or immunologic detection techniques. It is already possible to measure the interaction of specific agents with cellular target molecules, for example, through adduct formation with proteins and nucleic acids, excretion levels of excised adducts, or markers of altered gene expression.[24] the task of identifying the effects of lifestyle and other environmental and host factors is obviously formidable. Biochemical epidemiology represents an innovative approach that may help elucidate further the causes of cancer and the actual mechanisms of carcinogenesis.

DETERMINING CAUSALITY

In interpreting epidemiologic findings, one is guided by the magnitude of the risk estimates, their statistical significance (likelihood of being due to chance), and the rigor of the study design to avoid various kinds of bias, including those related to selection, confounding, classification, and measurement. A determination of causality in epidemiology is bolstered by dose-response relationships, the consistency and reproducibility of results, the strength and specificity of the association, its biological pausibility, and other considerations. Thus, inferences from epidemiology, as from other methods of inquiry, are not made in isolation but should take into account all relevant biologic information. Although epidemiologic and other observations can accumulate to the point that a causal hypothesis is likely, it is not possible to ever prove causality (in the strict sense, a hypothesis can only be disproved). Nevertheless, a causal hypothesis can be sufficiently probable, as in the case of menopausal estrogens and endometrial cancer and DES and vaginal adenocarcinoma, to provide a reasonable and even compelling basis for preventive and public health action.

REFERENCES

1. Hippocrates: On Airs, Waters, and Places. (Transl.) Classics 3:19, 1938
2. Rigoni-Stern D: Fatti statistici alle mallattie cancerose che servirono di base alle poche cose dette dal dott. G Progr Patol Terap 2nd Ser. 2:499, 1842; see also (transl) Scotto J, Bailar JC III: Rigoni-Stern and Medical Statistics: A nineteenth-century approach to cancer research. J Hist Med 65, 1969
3. Annonymous: Human papillomaviruses and cervical cancer. A fresh look at the evidence. Lancet 1:725, 1987
4. Gusberg SB: Precursors of corpus carcinoma, estrogens and adenomatous hyperplasia. Am J Obstet Gynecol 54:905, 1947
5. Brinton LA, Huggins GR, Lehman HF, et al: Long-term use of oral contraceptives and risk of invasive cervical cancer. Int. J Cancer 38:339, 1986
6. Hoover RN: Sex hormones and human carcinogenesis: Epidemiology. In Becker KL (ed): Principles and Practice of Endocrinology and Metabolism. JB Lippincott, Philadelphia, 1987
7. Hartge P, Hoover R, Lesher LP, et al: Talc and ovarian cancer. JAMA 250:1844, 1983
8. MacMahon B, Pugh TF: Epidemiology: Principles and Methods. Little, Brown, Boston, 1970
9. Doll R: The epidemiology of cancer. Cancer 45:2475, 1980
10. Fraumeni JF, Jr: Epidemiological approaches to cancer etiology. Annu Rev Public Health 3:85, 1982
11. Lilienfeld A, Pederson E, Dowd JE: Cancer Epidemiology: Methods of Study. Johns Hopkins Press, Baltimore, 1967
12. Cramer DW, Cutler SJ, Christine B: Trends in the incidence of endometrial cancer in the United States. Gynecol Oncol 2:130, 1974
13. Austin DF, Roe KM: The decreasing incidence of endometrial cancer: Public health implications. Am J Public Health 72:65, 1982
14. Cole P: Introduction: The analysis of case-control studies. p. 14. In Breslow NE, Day NE (eds): Statistical

Methods in Cancer Research. Vol. 1. International Agency for Research on Cancer, Lyons, 1980
15. Hoover RN: Hormonal, infectious and nutritional aspects of cancer of the female reproductive tract. p. 313. In Harris CC (ed): Biochemical and Molecular Epidemiology of Cancer. Alan R. Liss, New York, 1986
16. Hutchison GB: The epidemiologic method. p. 3. In Schottenfeld D, Fraumeni JF Jr (eds): Cancer Epidemiology and Prevention. WB Saunders, Philadelphia, 1982
17. Young JL Jr, Percy CL, Asire AJ (eds): Surveillance, Epidemiology, and End Results. National Cancer Institute Monograph 57. National Institutes of Health, US Department of Health and Human Services, Bethesda, MD, 1981
18. Hill AB: Principles of Medical Statistics. Oxford University Press, New York, 1966
19. Greene MH, Harris EL, Gershenson DM, et al: Melphalan may be a more potent leukemogen than cyclophosphamide. Ann Intern Med 105:360, 1986
20. Rothman KJ, Fyler DC, Goldblatt A, et al: Exogenous hormones and other drug exposures of children with congenital heart disease. Am J Epidemiol 109:433, 1979
21. Doll R, Peto R: The causes of cancer. J Natl Cancer Inst 66:1191, 1981
22. MacLure RM, MacMahon B: An epidemiologic perspective of environmental carcinogenesis. Epidemiol Rev 2:19, 1980
23. Tomatis L, Breslow NE, Bartsch H: Experimental studies in the assessment of cancer risk. p. 44. In Schottenfeld D, Fraumeni JF Jr (eds): Cancer Epidemiology and Prevention. WB Saunders, Philadelphia, 1982
24. Perera FP, Weinstein IB: Molecular epidemiology and carcinogen-DNA adduct detection: New approaches to studies of human cancer causation. J Chron Dis 35:581, 1982